

Claims

What is claimed is:

1. A system for extracting information from network data, comprising:  
an input interface connected to at least one source of network data; and  
5 a network event sensor, communicating with the input interface, the network event sensor applying at least a lexical engine to the network data to identify at least one network event.
2. The system of claim 1, wherein the at least one source of network data comprises an observation port connected to a network and continuously  
10 capturing network data from the network.
3. The system of claim 2, wherein the observation port comprises a network interface card.
4. The system of claim 3, wherein the network comprises at least one of an Ethernet network, a token ring network, and a TCP/IP network.
- 15 5. The system of claim 3, wherein the network interface card is invisible to the network.
6. The system of claim 1, wherein the at least one source of network data comprises stored network data.
7. The system of claim 6, wherein the stored network data comprise at least  
20 one of captured network files, Website mirrors, archives of Usenet files, and archives of email files.

09552878-042000

8. The system of claim 1, further comprising an interpreter module, the interpreter module scanning the network data to generate logical groupings of the network data.
9. The system of claim 8, wherein the logical groupings comprise packets.
- 5 10. The system of claim 8, wherein the interpreter module removes low-level encoding information from the network data to generate the logical groupings.
11. The system of claim 10, wherein the low-level encoding information removed by the interpreter module comprises hardware addressing information.
- 10 12. The system of claim 8, further comprising an assembler module, communicating with the interpreter module, the assembler module scanning the logical groupings to generate at least one session object.
13. The system of claim 12, wherein the at least one session object comprises at least one session file.
- 15 14. The system of claim 12, wherein the assembler module scans the logical groupings by examining at least one of source address, destination address, sequence numbers, source port, and destination port to generate the at least one session object.
15. The system of claim 12, wherein the network event sensor applies the
- 20 lexical engine to the at least one session object to identify the at least one network event as at least one of a predetermined set of event types.

000240" 8/82560

16. The system of claim 15, wherein the lexical engine detects the presence of at least one predefined keyword to identify the at least one of a predetermined set of event types.

17. The system of claim 16, wherein the predetermined set of event types  
5 comprises at least one of TCP, IP, UDP, SMTP, HTTP, NNTP, FTP, TELNET,  
DNS, RIP, BGP, MAIL, NEWS, HTML, XML, PGP, S/MIME, POP, IMAP,  
V-CARD, ICMP, NetBUI, IPX and SPX.

18. The system of claim 16, wherein the lexical engine accumulates a total  
number of occurrences for the at least one predefined keyword to identify the  
10 event type.

19. The system of claim 18, wherein the lexical engine applies a threshold to the number of occurrences to identify the event type.

20. The system of claim 12, wherein the network event sensor applies the  
lexical engine recursively to identify more than one event type contained in the  
15 at least one session object.

21. The system of claim 15, further comprising an extractor module, the extractor module extracting the at least one network event from the at least one session object according to the at least one of a predetermined set of event types.

20 22. The system of claim 21, wherein the extractor module comprises a library of extractor types, each of the extractor types corresponding to at least one of the at least one of a predetermined set of event types.

24. The system of claim 23, wherein the minimum subset of the network data is stored in a database.

26. The system of claim 1, wherein the network event sensor also applies a  
port detection engine to the network data to identify the at least one network  
10 event.

27. The system of claim 1, wherein the at least one source of network data comprises a plurality of sources of network data.

28. A method for extracting information from network data, comprising the steps of:

15           a) receiving network data from at least one source of network data; and

             b) applying at least a lexical engine to the network data to identify at

least one network event.

29. The method of claim 28, wherein the at least one source of network data  
comprises an observation port connected to a network and continuously  
20 capturing network data from the network.

30. The method of claim 29, wherein the observation port comprises a network interface card.

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

40. The method of claim 39, wherein the at least one session object comprises at least one session file.

1. **Introduction**  
 2. **Background**  
 3. **Methodology**  
 4. **Results**  
 5. **Discussion**  
 6. **Conclusion**  
 7. **References**  
 8. **Appendix**  
 9. **Figure 1**  
 10. **Figure 2**  
 11. **Figure 3**  
 12. **Figure 4**  
 13. **Figure 5**  
 14. **Figure 6**  
 15. **Figure 7**  
 16. **Figure 8**  
 17. **Figure 9**  
 18. **Figure 10**  
 19. **Figure 11**  
 20. **Figure 12**  
 21. **Figure 13**  
 22. **Figure 14**  
 23. **Figure 15**  
 24. **Figure 16**  
 25. **Figure 17**  
 26. **Figure 18**  
 27. **Figure 19**  
 28. **Figure 20**  
 29. **Figure 21**  
 30. **Figure 22**  
 31. **Figure 23**  
 32. **Figure 24**  
 33. **Figure 25**  
 34. **Figure 26**  
 35. **Figure 27**  
 36. **Figure 28**  
 37. **Figure 29**  
 38. **Figure 30**  
 39. **Figure 31**  
 40. **Figure 32**  
 41. **Figure 33**  
 42. **Figure 34**  
 43. **Figure 35**  
 44. **Figure 36**  
 45. **Figure 37**  
 46. **Figure 38**  
 47. **Figure 39**  
 48. **Figure 40**  
 49. **Figure 41**  
 50. **Figure 42**  
 51. **Figure 43**  
 52. **Figure 44**  
 53. **Figure 45**  
 54. **Figure 46**  
 55. **Figure 47**  
 56. **Figure 48**  
 57. **Figure 49**  
 58. **Figure 50**  
 59. **Figure 51**  
 60. **Figure 52**  
 61. **Figure 53**  
 62. **Figure 54**  
 63. **Figure 55**  
 64. **Figure 56**  
 65. **Figure 57**  
 66. **Figure 58**  
 67. **Figure 59**  
 68. **Figure 60**  
 69. **Figure 61**  
 70. **Figure 62**  
 71. **Figure 63**  
 72. **Figure 64**  
 73. **Figure 65**  
 74. **Figure 66**  
 75. **Figure 67**  
 76. **Figure 68**  
 77. **Figure 69**  
 78. **Figure 70**  
 79. **Figure 71**  
 80. **Figure 72**  
 81. **Figure 73**  
 82. **Figure 74**  
 83. **Figure 75**  
 84. **Figure 76**  
 85. **Figure 77**  
 86. **Figure 78**  
 87. **Figure 79**  
 88. **Figure 80**  
 89. **Figure 81**  
 90. **Figure 82**  
 91. **Figure 83**  
 92. **Figure 84**  
 93. **Figure 85**  
 94. **Figure 86**  
 95. **Figure 87**  
 96. **Figure 88**  
 97. **Figure 89**  
 98. **Figure 90**  
 99. **Figure 91**  
 100. **Figure 92**  
 101. **Figure 93**  
 102. **Figure 94**  
 103. **Figure 95**  
 104. **Figure 96**  
 105. **Figure 97**  
 106. **Figure 98**  
 107. **Figure 99**  
 108. **Figure 100**  
 109. **Figure 101**  
 110. **Figure 102**  
 111. **Figure 103**  
 112. **Figure 104**  
 113. **Figure 105**  
 114. **Figure 106**  
 115. **Figure 107**  
 116. **Figure 108**  
 117. **Figure 109**  
 118. **Figure 110**  
 119. **Figure 111**  
 120. **Figure 112**  
 121. **Figure 113**  
 122. **Figure 114**  
 123. **Figure 115**  
 124. **Figure 116**  
 125. **Figure 117**  
 126. **Figure 118**  
 127. **Figure 119**  
 128. **Figure 120**  
 129. **Figure 121**  
 130. **Figure 122**  
 131. **Figure 123**  
 132. **Figure 124**  
 133. **Figure 125**  
 134. **Figure 126**  
 135. **Figure 127**  
 136. **Figure 128**  
 137. **Figure 129**  
 138. **Figure 130**  
 139. **Figure 131**  
 140. **Figure 132**  
 141. **Figure 133**  
 142. **Figure 134**  
 143. **Figure 135**  
 144. **Figure 136**  
 145. **Figure 137**  
 146. **Figure 138**  
 147. **Figure 139**  
 148. **Figure 140**  
 149. **Figure 141**  
 150. **Figure 142**  
 151. **Figure 143**  
 152. **Figure 144**  
 153. **Figure 145**  
 154. **Figure 146**  
 155. **Figure 147**  
 156. **Figure 148**  
 157. **Figure 149**  
 158. **Figure 150**  
 159. **Figure 151**  
 160. **Figure 152**  
 161. **Figure 153**  
 162. **Figure 154**  
 163. **Figure 155**  
 164. **Figure 156**  
 165. **Figure 157**  
 166. **Figure 158**  
 167. **Figure 159**  
 168. **Figure 160**  
 169. **Figure 161**  
 170. **Figure 162**  
 171. **Figure 163**  
 172. **Figure 164**  
 173. **Figure 165**  
 174. **Figure 166**  
 175. **Figure 167**  
 176. **Figure 168**  
 177. **Figure 169**  
 178. **Figure 170**  
 179. **Figure 171**  
 180. **Figure 172**  
 181. **Figure 173**  
 182. **Figure 174**  
 183. **Figure 175**  
 184. **Figure 176**  
 185. **Figure 177**  
 186. **Figure 178**  
 187. **Figure 179**  
 188. **Figure 180**  
 189. **Figure 181**  
 190. **Figure 182**  
 191. **Figure 183**  
 192. **Figure 184**  
 193. **Figure 185**  
 194. **Figure 186**  
 195. **Figure 187**  
 196. **Figure 188**  
 197. **Figure 189**  
 198. **Figure 190**  
 199. **Figure 191**  
 200. **Figure 192**  
 201. **Figure 193**  
 202. **Figure 194**  
 203. **Figure 195**  
 204. **Figure 196**  
 205. **Figure 197**  
 206. **Figure 198**  
 207. **Figure 199**  
 208. **Figure 200**  
 209. **Figure 201**  
 210. **Figure 202**  
 211. **Figure 203**  
 212. **Figure 204**  
 213. **Figure 205**  
 214. **Figure 206**  
 215. **Figure 207**  
 216. **Figure 208**  
 217. **Figure 209**

41. The method of claim 39, wherein the step (e) of scanning the logical groupings comprises a step of f) examining at least one of source address, destination address, sequence numbers, source port, and destination port to generate the at least one session object.

5 42. The method of claim 39, further comprising a step of g) identifying the at least one network event as at least one of a predetermined set of event types.

43. The method of claim 42, wherein the step (g) of identifying comprises a step of (h) detecting the presence of at least one predefined keyword to identify the at least one of a predetermined set of event types.

10 44. The method of claim 43, wherein the predetermined set of event types comprises at least one of TCP, IP, UDP, SMTP, HTTP, NNTP, FTP, TELNET, DNS, RIP, BGP, MAIL, NEWS, HTML, XML, PGP, S/MIME, POP, IMAP, V-CARD, ICMP, NetBUI, IPX and SPX.

15 45. The method of claim 43, wherein the step (h) of detecting comprises a step of (i) accumulating a total number of occurrences for the at least one predefined keyword to identify the event type.

46. The method of claim 45, wherein the step (h) of detecting comprises a step (j) of applying a threshold to the number of occurrences to identify the event type.

20 47. The method of claim 39, wherein the step of b) applying at least the lexical engine comprises a step of k) applying the lexical engine recursively to identify more than one event type contained in the at least one session object.

000240" 8282560

48. The method of claim 42, further comprising a step of l) extracting the at least one network event from the at least one session object according to the at least one of a predetermined set of event types.

49. The method of claim 48, wherein the step (l) of extracting comprises a  
5 step of m) selecting at least one extractor module from a library of extractor types, each of the extractor types corresponding to at least one of the at least one of a predetermined set of event types.

50. The method of claim 49, further comprising a step of n) storing a minimum subset of the network data to reconstruct the at least one network  
10 event.

51. The method of claim 50, wherein the step (n) of storing comprises a step o) of storing the minimum subset of the network data in a database.

52. The method of claim 51, further comprising a step of p) querying the database for information related to the at least one network event.

15 53. The method of claim 28, further comprising a step q) of applying a port detection engine to the network data to identify the at least one network event.

54. The method of claim 28, wherein the at least one source of network data comprises a plurality of sources of network data.

09552878 042000